

1 **ABSTRACT**

2 Potentially identical objects (e.g., files) are located across multiple
3 computers based on stochastic partitioning of workload. For each of a plurality of
4 objects stored on a plurality of computers in a network, a portion of object
5 information corresponding to the object is selected. The object information can be
6 generated in a variety of manners (e.g., based on hashing the object, based on
7 characteristics of the object, and so forth). Any of a variety of portions of the
8 object information can be used (e.g., the least significant bits of the object
9 information). A stochastic partitioning process is then used to identify which of
10 the plurality of computers to communicate the object information to for
11 identification of potentially identical objects on the plurality of computers.
12
13
14
15
16
17
18
19
20
21
22
23
24
25